State Amplification Subject To Masking Constraints

O. Ozan Koyluoglu, Rajiv Soundararajan, and Sriram Vishwanath

Abstract

This paper considers a state dependent channel with one transmitter, Alice, and two receivers, Bob and Eve. The problem is to effectively convey ("amplify") the channel state sequence to Bob while "masking" it from Eve. The extent to which the state sequence cannot be masked from Eve is referred to as leakage. This can be viewed as a secrecy problem, where we desire that the channel state itself be minimally leaked to Eve while being communicated to Bob. The paper is aimed at characterizing the trade-off region between amplification and leakage rates for such a system.

An achievable coding scheme is presented, wherein the transmitter transmits a partial state information over the channel to facilitate the amplification process. For the case when Bob observes a stronger signal than Eve, the achievable coding scheme is enhanced with *secure refinement*.

Outer bounds on the trade-off region are also derived, and used in characterizing some special case results. In particular, the optimal amplification-leakage rate difference, called as differential amplification capacity, is characterized for the reversely degraded discrete memoryless channel, the degraded binary, and the degraded Gaussian channels. In addition, for the degraded Gaussian model, the extremal corner points of the trade-off region are characterized, and the gap between the outer bound and achievable rate-regions is shown to be less than half a bit for a wide set of channel parameters.

I. Introduction

Information-theoretic secrecy analysis is now a very well established field (see, e.g., Special Issue on Information Theoretic Security, *IEEE Trans. Inf. Theory*, June 2008 and references therein). A vast body of this literature is focused on secrecy metrics and information-theoretic secrecy guarantees for messages from passive and active adversaries. Although information-theoretic secrecy is of considerable interest, there is still a disconnect between the domain of information-theoretic secrecy and that of computational secrecy methodologies in cryptography. This disconnect, along with conservative rate guarantees, have meant that information-theoretic secrecy has found only a limited number of applications.

In recent years, a bridge between cryptography and information theory has emerged in the form of channel state-dependent key generation [1], [2], [3], [4]. In this line of work, a state dependent channel (such as the wireless channel) is considered, and a function of this state is intended to be a 'shared secret' between the legitimate transmitter (Alice) and legitimate receiver (Bob) while aiming to keep the eavesdropper (Eve) as much in the dark as possible. We intend to provide information-theoretic guarantees to the extent to which such a shared secret can be realized for our system model. Once obtained, this shared secret can now be employed to seed numerous symmetric key cryptosystems [5], [6]. Another relevant setting where our results will be of significant interest is in cognitive radio systems [7], [8], [9], [10], [11], [12], [13]. In this scenario, the cognitive encoder (Alice) facilitates the secure communication of the primary signal (the state sequence of the channel) by amplifying the signal at the primary receiver (Bob) while masking it from the eavesdropper (Eve). More generally, this setting belongs to user cooperation or relaying architectures that increase/decrease the security of the communication systems [14], [15], [16], [17].

In this paper, we consider a state dependent broadcast channel model with two users, and consider the question of to what extent the state of the channel can be amplified at the receiver (Bob) and masked from the other receiver (called as Eve). In the best case, the state(s) seen by Bob and Eve will be completely different (independent). However, we consider what might be a pessimistic model where there is a single channel state defining the channel for both Bob and Eve. Moreover, the entire channel state sequence is presumed to be known non-causally to the transmitter (a Gel'fand-Pinsker-style assumption [18]). The only manner in which an asymmetry can be affected between Bob and Eve is by the encoding used at the transmitter. For such a system, we aim to characterize the trade-off between the "amplification"-rate at which the legitimate pair can operate and the "leakage"-rate to the eavesdropper. In essence, as long as there is a non-trivial difference between the two, this can be used to develop shared keys and enable cryptographic algorithms. More specifically, we define the measure differential amplification capacity as the maximal knowledge difference regarding the state of the channel between Bob and Eve. This metric transforms the two-dimensional amplification-leakage trade-off to a single dimension, and can be viewed as a bound on the secrecy that can be leveraged from the channel state knowledge difference. In this form, the problem at hand is very much related to the generation of secrecy using sources and channels. Here, not only the source in the model is the channel state (different from the models studied in [19], [20]), but also it is non-trivially combined with the encoded signals of Alice (via the state dependent channel) to produce the observations at Bob and Eve. Hence, our formulation can be closely associated with the problem of secret key agreement over state dependent channels [21], [22]. In such problems, one is interested in the

The authors are with the Wireless Networking and Communications Group, Department of Electrical and Computer Engineering, The University of Texas, Austin, TX 78712. Email: ozan@austin.utexas.edu, rajivs@utexas.edu, sriram@austin.utexas.edu.

The material in this paper was presented in part at the Forty-Ninth Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL in September 2011.

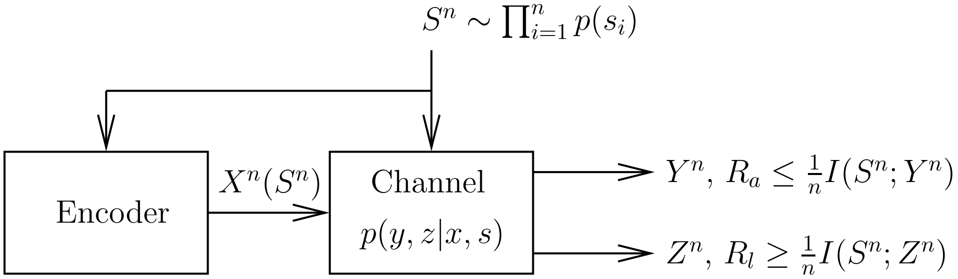
$S^n \sim \prod_{i=1}^n p(s_i)$
Encoder
$X^n(S^n)$
Channel
$p(y,z|x,s)$
$Y^n,\ R_a \leq \frac{1}{n}I(S^n;Y^n)$
$Z^n,\ R_l \geq \frac{1}{n}I(S^n;Z^n)$

Fig. 1. The system model for amplification subject to masking problem.

design of coding strategies that allow an agreement of secure bits between the legitimate users utilizing the state dependent channel. For example, sending secure bits over the channel will increase the secret key rate [22], [23], [24]. On the other hand, the problem studied in this work, when specialized to the differential amplification metric, considers only the channel state knowledge difference and provides a bound on the secret bits that can be extracted only from the channel state ([1], [2], [3], [4]). In general, we are not only interested in the knowledge difference, but also in the entire rate trade-off region, which constitutes the main focus of this paper.

A. Problem Statement

Consider a discrete memoryless channel given by $p(y,z|x,s)$, where $x \in \mathcal{X}$ is the channel input, $s \in \mathcal{S}$ is the channel state, and $(y,z) \in (\mathcal{Y} \times \mathcal{Z})$ is the channel output, with y corresponding to the legitimate receiver (Bob) and z the eavesdropper (Eve). The channel is memoryless in the sense that

$$p(Y^n = y^n, Z^n = z^n | X^n = x^n, S^n = s^n) = \prod_{i=1}^{n} p(y(i), z(i) | x(i), s(i)), \tag{1}$$

and the state sequence is independent and identically distributed (i.i.d.) according to a probability distribution indicated by $p(s)$. It is assumed that the channel state sequence is non-causally known at the transmitter. (The system model is depicted in Fig 1.)

The task of the encoder is to “amplify” the state sequence at Bob (channel output Y^n) and to “mask” the state sequence from Eve (Z^n). Formally, we measure the former by the state amplification rate R_a and the latter as the state leakage rate R_l. We say (R_a, R_l) is achievable, if for any given $\epsilon > 0$,

$$\frac{1}{n} I(S^n; Y^n) \geq R_a - \epsilon \tag{2}$$

$$\frac{1}{n} I(S^n; Z^n) \leq R_l + \epsilon \tag{3}$$

for sufficiently large n. The problem is to characterize all achievable (R_a,R_l) pairs, which we denote by the trade-off region $\mathcal{C}$.

The performance of the encoder is also quantified by measuring the difference between the achievable amplification and leakage rates. The *differential amplification rate* R_d is said to be achievable if $R_d = R_a - R_l$ for some $(R_a, R_l) \in \mathcal{C}$. The maximum value of the differential amplification rate is called as the *differential amplification capacity,* denoted by C_d, where

$$C_d = \sup_{(R_a, R_l) \in \mathcal{C}} R_a - R_l. \tag{4}$$

Note that this quantity is important for applications like key agreement from the state of the channel, as this difference measures the knowledge difference between the two receivers regarding the state of the channel.

A cost constraint may also be imposed on the channel input with

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\{c(X(i))\} \leq C, \tag{5}$$

where $c : \mathcal{X} \to \mathbb{R}^+$ defines the cost per input letter and the expectation is over the distribution of the channel input. In this scenario, we say (R_a, R_l) is achievable under the cost function $c(.)$ and cost C, if (2), (3), and (5) are satisfied in the limit of large n. (We use this constraint for the Gaussian channel, where the cost is the average transmitted power.)

Remark: One can say that the equivocation rate Δ_l is achievable, if the signaling scheme satisfies

$$\Delta_l \leq \frac{1}{n} H(S^n | Z^n) + \epsilon, \tag{6}$$

and, correspondingly consider the achievable (R_a, Δ_l) pairs. Hence, the problem can be re-stated in terms of equivocation rate, where we seek to characterize all achievable (R_a, Δ_l) pairs in the limit of large n. Since both the equivocation and leakage rate notions characterize the same trade-off, both notions can be used interchangeably.

B. Related work, summary of results, and organization

The problem of communication over state dependent channels is studied by Gel'fand and Pinsker [18], where a message has to be reliably transmitted over the channel with non-causal state knowledge at the transmitter. The Gaussian version of the problem is solved in [25] through the famous dirty paper coding scheme. While the wiretap channel is introduced and solved in [26], these results are extended to a broadcast setting in [27]. The problem of sending secure messages over state dependent wiretap channels is studied in [23], [24].

On the other hand, the problems of state amplification and state masking are individually solved in [28], [29], [30] for point-to-point channels. Both [28], [29] and [30] consider the problem of reliable transmission of messages in addition to state amplification and state masking respectively. In this paper, we consider the problem of amplifying the state to a desired receiver while trying to minimize the leakage (or mask the state) to the eavesdropper.

We note that, if we set $R_a = 0$ in our problem definition, it reduces to the state masking problem as studied in [30]. In other words,

$$\begin{aligned} R_a &= 0 & (7)\\ R_l &= \min_{p(x|s) \text{ s.t. } \mathbb{E}\{c(X)\} \leq C} I(S;Z) & (8) \end{aligned}$$

can be shown to be achievable [30]. Also, when $R_l \geq H(S)$, the problem reduces to a state amplification problem [29], and one can achieve the following rate pair.

$$\begin{aligned} R_a &= \min\{H(S), \max_{p(x|s)} I(X,S;Y)\} & (9)\\ R_l &\geq H(S) & (10) \end{aligned}$$

These represent two extremes of the trade-off region between the amplification and masking rates. In this paper, we aim at developing an understanding of this trade-off region through achievable regions and outer bounds, and characterizing special cases when they match. Our main results can be summarized as:

- **Achievable Regions:** Our achievability arguments are based on enumerating typical state sequences using two indices, and sending one of the indices over the channel. Towards this end, we construct a codebook corresponding to the codeword carrying this index (denoted by U^n) in such a way that reliable communication can be achieved over the state-dependent channel. Subsequently, we derive expressions for achievable amplification and leakage rates by determining single-letter bounds on $\frac{1}{n}I(S^n;Y^n)$ and $\frac{1}{n}I(S^n;Z^n)$ respectively, and the achievable region is established over the input probability distributions $p(u,x|s)$. The bounds show that the rate of the message index not only amplifies the R_a, but also increases the leakage R_l, thereby establishing a trade-off for implementation.
- **Secure Refinement:** We also show that it is possible to extend the proposed region with *secure refinement* when Bob observes a "stronger" channel than Eve. In precise terms, this corresponds to instances of $p(u,x|s)$ satisfying $I(U;Y) \geq I(U;Z)$. Note that a channel is said to have a less noisy structure if $I(U;Y) \geq I(U;Z)$ for all input probability distributions [31], [32]. We find that the utilization of the notion of secure refinement approach is critical to such channels, and we show that the leakage due to transmission of the message can be minimized by securing (a part of) the message.
- **Special Classes of Channels:** Our outer bound arguments are based on upper bounding $\frac{1}{n}I(S^n;Y^n)$ and lower bounding $\frac{1}{n}I(S^n;Z^n)$. The achievable schemes and outer bounds presented are used to establish optimality results for a class of channels. In particular, we show that the proposed scheme achieves the optimal differential amplification capacity (i.e., the maximum value of $R_a - R_l$ over the set of achievable (R_a, R_l) pairs) for the reversely degraded discrete memoryless channel (DMC), the degraded binary channel, and the degraded Gaussian channel.
- **Gaussian Channels:** We characterize the corner points of the region for the degraded Gaussian channel. In this scenario, we further bound the gap between achievable and converse regions, and show the following: Let us denote the message capacity of Bob's channel as $C_b = \frac{1}{2}\log(1+\text{SNR}_b)$ and that of Eve's channel as $C_e = \frac{1}{2}\log(1+\text{SNR}_e)$, where SNR_b (SNR_e) is the signal-to-noise ratio of Bob (respectively, Eve). Then, for any given leakage rate R_l, the gap between the upper and lower bounds on the amplification rate R_a is bounded by C_b. Similarly, for any given amplification rate R_a, the achievable leakage rate is within C_e of the lower bound on R_l. In particular, the corresponding gaps are within half a bit when $\text{SNR}_b \leq 1$ and $\text{SNR}_e \leq 1$ respectively.

The rest of the paper is organized as follows. Section II presents our main results, where we provide our proposed coding schemes and outer bounding arguments. Section III provides optimality discussions and numerical results for special classes of DMCs including the reversely degraded channel, modulo additive binary channel model, and the memory with defective cells model. The Gaussian channel model is considered in Section IV along with corresponding optimality results. Finally, we conclude the paper in Section V. The proofs are collected in Appendices to improve the flow of the paper.

II. Main Results

A. Achievable Regions

We have the following propositions for any given $p(s)$ and the channel $p(y,z|x,s)$.

1) State sequence covering: The scheme presented below is based on communicating a covering of the state sequence while ensuring that the covering is decodable at Bob. We have the following result.

Proposition 1: Let $\mathcal{R}^1$ be the closure of the union of all (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq I(S;Y,U) \\ R_l &\geq \min\{I(S;Z,U), I(U,S;Z)\} \\ 0 &\leq I(U;Y) - I(U;S), \end{aligned}$$

over all distributions $p(u,x|s)$. Then, $\mathcal{R}^1 \subseteq \mathcal{C}$.

Proof: See Appendix A. ■

We note that, provided $I(U;Y) - I(U;S) \geq 0$, the covering codeword can be decoded at Bob. Then, the state uncertainty can be reduced from $H(S)$ to $H(S|Y,U)$ by listing S^n sequences that are jointly typical with (U^n, Y^n). This will give rise to the expression $I(S;Y,U)$, as presented in the proposition. (Leakage expressions follow by a similar argument, where U^n sequence is added to the expression $\frac{1}{n}I(S^n;Z^n)$ to derive the achievable leakage expressions.)

2) State enhanced messaging: In order to achieve a better rate region, the encoder can send a message over the state dependent channel, where the message carries partial information about the state sequence. The corresponding achievable region is given by the following result.

Proposition 2: Let $\mathcal{R}^2$ be the closure of the union of all (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq \min\{H(S), I(U,S;Y)\} \\ R_l &\geq I(U,S;Z) \\ 0 &\leq I(U;Y) - I(U;S), \end{aligned}$$

over all distributions $p(u,x|s)$. Then, $\mathcal{R}^2 \subseteq \mathcal{C}$.

Proof: See Appendix B. ■

The achievable rate region can be interpreted as follows: The rate $I(U;Y) - I(U;S)$ corresponds to the Gel'fand-Pinsker message rate that can be reliably communicated over the channel. As long as this rate is positive for a given input probability distribution, the codeword U^n can be reliably communicated over the channel. Bob can decode U^n from Y^n by employing a jointly-typical decoder. Subsequently, the state uncertainty can be reduced from $H(S)$ to $H(S|Y,U)$ by listing S^n sequences that are jointly typical with (U^n, Y^n). Further, the rate $I(U;Y) - I(U;S)$ provides an additional refinement to the uncertainty, which together with $I(S;Y,U)$ sums to $I(U,S;Y)$. (See also [29], where the authors show that it is possible to interpret this scheme as a source coding method with Wyner-Ziv coding [33].) The analysis of R_l follows by analyzing an upper-bound on the leakage rate by adding the codeword sequence to the expression, $\frac{1}{n}I(S^n;Z^n) \leq \frac{1}{n}I(U^n,S^n;Z^n)$, which can be single-letterized as given in Appendix B.

We observe that the leakage expression can be enhanced when $I(U;Z) \geq I(U;Y)$, and this is detailed in the following proposition.

Proposition 3: For all input distributions $p(u,x|s)$ that satisfy $I(U;Z) \geq I(U;Y)$. R_l in Proposition can be bounded as

$$R_l \geq I(S;Z,U) + R_u,$$

for some $R_u \leq \min\{I(U;Y) - I(U;S), H(S|U,Y)\}$, while $R_a \leq I(S;Y,U) + R_u$. In particular, for $R_u = \min\{I(U;Y) - I(U;S), H(S|U,Y)\}$, $R_a \leq \min\{H(S), I(U,S;Y)\}$, and

$$\begin{aligned} R_l \geq \ & \min\{I(S;Z,U) + I(U;Y) - I(U;S) = I(U,S;Z) - [I(U;Z) - I(U;Y)], \\ & H(S) - [H(S|Z,U) - H(S|Y,U)]\}. \end{aligned}$$

Thus, Proposition 3 can be enhanced for those input distributions satisfying $I(U;Z) \geq I(U;Y)$.

Proof: (Sketch of the proof) The requirement of $I(U;Z) \geq I(U;Y)$ enables the decodability of U^n at Eve. Then, using arguments similar to those used for the amplification bound in the proof of Proposition 3 (and upper bounding the inequalities from (66) to (74) by reversing the ones given in (75) to (78)), we obtain the R_l bound as

$$R_l \geq I(S;Z,U) + R_u. \tag{11}$$

Further, by choosing $R_u \leq \min\{I(U;Y) - I(U;S), H(S|U,Y)\}$ we have the desired result. ■

The three results above (Propositions 1, 3, and 3) is combined in the following theorem.

Theorem 4: Let $\mathcal{R}^3$ be the closure of the union of all (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq I(S;Y,U) + R_u \\ R_l &\geq \min\left\{I(U,S;Z), I(S;Z,U) + R_u\right\} \\ R_u &\leq \min\{I(U;Y) - I(U;S), H(S|Y,U)\}, \end{aligned}$$

over all distributions $p(u,x|s)$ satisfying $I(U;Y) \geq I(U;S)$. Then, $\mathcal{R}^3 \subseteq \mathcal{C}$.

Remarkably, this representation also shows that the covering strategy is a special case of the state enhanced messaging approach (when $R_u = 0$ is chosen). We also note that, in the region above, increasing R_u will not only increase the amplification rate but will also increase the leakage rate. Thus, for some scenarios, implementing only the covering scheme might be advantageous. Further implications of this observation on the amplification-leakage region is discussed in Section III-A.

3) Secure refinement: Consider all input distributions $p(u,x|s)$ satisfying $I(U;Y) \geq I(U;Z)$. For such distributions, it is possible to send refinement information securely over the channel. This way, the leakage increase due to refinement index is decreased as the security of the index will lower the corresponding leakage rate achieved at Eve compared to the non-secured case.

Theorem 5: Let $\mathcal{R}^4$ be the closure of the union of all (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq I(S;Y,U) + R_u \\ R_l &\geq \min\left\{I(U,S;Z), I(S;Z,U) + \min\{R_u, [I(U;Z) - I(U;S)]^+\}\right\} \\ R_u &\leq \min\{I(U;Y) - I(U;S), H(S|Y,U)\} \end{aligned}$$

over all distributions $p(u,x|s)$ satisfying $I(U;Y) \geq I(U;S)$ and $I(U;Y) \geq I(U;Z)$. Then, $\mathcal{R}^4 \subseteq \mathcal{C}$.

Proof: See Appendix C. ■

Comparing this region with that presented in Proposition 3 and Theorem 4, we observe that the leakage rate is decreased when the refinement is secured from Eve.

Note that, when $I(U;Z) \geq I(U;S)$, the analysis given in Appendix C for the region above, counts the part of the message with rate $[I(U;Z) - I(U;S)]^+$ as leakage (as this part is not secured and still being used as a message relevant to state). In other words, one can view the codewords as $U^n(W_u, W_u', W_u'')$ with rates $R_u = I(U;Y) - I(U;Z)$, $R_u' = I(U;Z) - I(U;S)$, $R_u'' = I(U;S)$, where W_u is secured refinement and W_u' is common refinement that is not only used at Bob, but also is leaked to Eve in the analysis. By modifying the coding scheme, one can utilize only the secure rate as the refinement (i.e., only W_u carries information regarding S^n). With such an approach, it is possible to enhance the leakage given in Theorem 5 (although at the cost of a reduction in the amplification rate) as given in the following proposition.

Proposition 6: The region of the union of all (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq I(S;Y,U) + R_u \\ R_l &\geq \min\{I(U,S;Z), I(S;Z,U)\} \\ R_u &\leq \min\{I(U;Y) - I(U;Z), H(S|Y,U)\} \end{aligned}$$

for the input distributions satisfying $I(U;Y) \geq I(U;Z) \geq I(U;S)$ is achievable.

Proof: (Sketch of the proof) Generate $U^n(W_u, W_u', W_u'')$ codewords with $R_u'' = I(U;S)$ and $R_u' = I(U;Z) - I(U;S)$ for the case of $I(U;Z) \geq I(U;S)$. W_u carries information regarding the state, W_u' is randomly chosen, and W_u'' is chosen to find an index l such that $u^n(w_u, w_u', l)$ is jointly typical with $s^n(w_u, w_r)$. Then, along the lines of the proof of Proposition 5, we obtain the result above. ■

Combining Proposition 6 with the case $I(U;Z) \leq I(U;S)$ of Theorem 5 (see Appendix C), we obtain the following.

Theorem 7: Let $\mathcal{R}^5$ be the closure of the union of all (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq I(S;Y,U) + R_u \\ R_l &\geq \min\{I(U,S;Z), I(S;Z,U)\} \\ R_u &\leq \min\left\{I(U;Y) - \max\{I(U;S), I(U;Z)\}, H(S|Y,U)\right\} \end{aligned}$$

over all distributions $p(u,x|s)$ satisfying $I(U;Y) \geq I(U;S)$ Then, $\mathcal{R}^5 \subseteq \mathcal{C}$.

Note that, the amplification rate that can be obtained with such an approach is lower than the previous case (Theorem 5), as the message rate $R_u \leq I(U;Y) - I(U;Z) < I(U;Y) - I(U;S)$ if $I(U;Z) > I(U;S)$. Therefore, the improvement on the leakage expression compared to Theorem 5 is obtained with a degradation on the amplification rate.

B. Outer Bounds

We now derive upper bounds on R_a and lower (upper) bounds on R_l (respectively, on Δ_l).

Proposition 8: If (R_a, R_l) is achievable, then $(R_a, R_l) \in \mathcal{R}_o^1$, where

$$\mathcal{R}_o^1 = \bigcup_{p(u,x|s)} (R_a, R_l)$$

satisfying

$$\begin{aligned} R_a &\leq \min\{H(S), I(X,S;Y)\} \\ R_l &\geq I(S;Z,U) \\ 0 &\leq I(U;Z) - I(U;S), \end{aligned}$$

for any given $p(u,x|s)$.

Proof: See Appendix D. ■

We now provide an outer bound for the degraded channel $p(y,z|x,s) = p(y|x,s)p(z|y)$ using the result above.

Proposition 9: If the channel satisfies $p(y,z|x,s) = p(y|x,s)p(z|y)$ and if (R_a, R_l) is achievable, then $(R_a, R_l) \in \mathcal{R}_o^2$, where

$$\mathcal{R}_o^2 = \bigcup_{p(u,x|s)} (R_a, R_l)$$

satisfying

$$\begin{aligned} R_a &\leq \min\{H(S), I(X,S;Y)\} \\ R_l &\geq I(S;Z,U) \\ R_a - R_l &\leq I(X,S;Y|Z) \\ 0 &\leq I(U;Y) - I(U;S), \end{aligned}$$

for any given $p(u,x|s)$.

Proof: See Appendix E. ■

III. Special discrete memoryless channel models

A. Reversely degraded DMCs

We say that the channel is reversely degraded if $(X,S) \to Z \to Y$ forms a Markov Chain. Note that, this corresponds to a stronger channel seen by Eve compared to that of Bob. We have the following result for this set of channels.

Theorem 10: The optimal differential amplification rate for reversely degraded DMCs is given by

$$C_d = \max_{p(x|s)} I(S;Y) - I(S;Z)$$

Proof: Achievability of the stated difference follows from Proposition 1 by substituting $U = \phi$. We provide the converse in Appendix F. ■

Note that coding can not improve this difference as the channel is reversely degraded. Thus, coding might help to increase R_a at the expense of decreasing $R_a - R_l$ for the reversely degraded scenario. In particular, the region provided by Proposition 3, which is an enhancement over that of Proposition 3, is achievable for the reversely degraded scenario. And, in the region stated in Proposition 3, R_a vs. $R_a - R_l$ can be traded-off using different input distributions. ($U = \phi$ case will correspond to the maximum $R_a - R_l$, and achieve C_d.)

B. Modulo additive binary model

Consider the channels given by

$$\begin{aligned} Y(i) &= X(i) \oplus S(i) \oplus N(i) \\ Z(i) &= X(i) \oplus S(i) \oplus N_z(i), \end{aligned} \tag{12}$$

where the state and noise distributions are generated i.i.d. as $S(i) \sim$ Bern(p_s), $N(i) \sim$ Bern(p_n), $N_z(i) \sim$ Bern(p_{n_z}). (All p_ks satisfy $p_k \in [0, 0.5]$ for $k = \{s, n, n_z\}$.) In this section, we use the following notation for the binary convolution $p \otimes q \triangleq p(1-q) + q(1-p)$.

1) State cancelation scheme: To cancel the state from the channel, we send

$$X(i) = U(i) \oplus S(i), \tag{13}$$

where $U(i) \sim \text{Bern}(p_u)$ and the codewords U^n carry a description of the state sequence S^n. This way, we achieve the following inner-bound.

Corollary 11: The state cancelation scheme, which sends Bern(p_u) distributed signal XORed with state sequence at each time instant, achieves the set of (R_a, R_l) pairs denoted by the region $\mathcal{R}^{\text{SC}}$, where

$$\mathcal{R}^{\text{SC}} = \text{Convex Hull}\left\{ \bigcup_{p_u \in [0,0.5], p_u \otimes p_s \leq 0.5} (R_a(p_u), R_l(p_u)) \right\} \subseteq \mathcal{C},$$

with

$$\begin{aligned} R_a(p_u) &\leq \min\left\{H(p_s), H(p_u \otimes p_n) - H(p_n)\right\} \\ R_l(p_u) &\geq H(p_u \otimes p_{n_z}) - H(p_{n_z}). \end{aligned}$$

Proof: Achievability follows from Proposition 3. ■

2) Optimal differential amplification rate:

Corollary 12: If $p_n \leq p_{n_z}$ and $H(p_s) \geq 1 - H(p_n)$ for a binary model the optimal amplification and leakage rate difference is given by

$$C_d = H(p_{n_z}) - H(p_n).$$

Proof: From Proposition 9, we obtain the following. If $p_n \leq p_{n_z}$, any given $(R_a, R_l) \in \mathcal{C}$ satisfies

$$R_a - R_l \leq H(p_{n_z}) - H(p_n) + \max_{p(x|s)} \left\{H(X \oplus S \oplus N) - H(X \oplus S \oplus N_z)\right\}. \tag{14}$$

Note that, this upper-bound can be evaluated by observing

$$\begin{aligned} &\max_{p(x|s)} \left\{H(X \oplus S \oplus N) - H(X \oplus S \oplus N_z)\right\} \\ &= \max_{p(x|s)} \left\{H(X \oplus S \oplus N) - H(X \oplus S \oplus N \oplus N_z^*)\right\} \\ &\leq \max_{p(x|s)} \left\{H(X \oplus S \oplus N) - H(X \oplus S \oplus N \oplus N_z^* | N_z^*)\right\} \\ &= 0, \end{aligned} \tag{15}$$

where the equality is due to the channel degradedness condition with appropriate noise term N_z^* independent of N such that $N \oplus N_z^* = N_z$, and the inequality is due to the fact that conditioning does not increase the entropy. Using this we observe that the outer-bound is maximized with a choice of $p(x) = 0.5$, which evaluates to

$$R_a - R_l \leq H(p_{n_z}) - H(p_n). \tag{16}$$

This expression is achieved by Corollary 11, when we choose $p(u) = 0.5$, if $H(p_s) \geq 1 - H(p_n)$.

■

C. Memory with defective cells model

We consider the model of information transmission over write-once memory device with stuck-at defective cells [34], [35]. In this channel model, each memory cell corresponds to a channel state instant with cardinality $|\mathcal{S}| = 3$, where the binary channel output is determined from the binary channel input and the channel state as:

$$\begin{aligned} p(y = 0 | x, s = 0) &= 1 \\ p(y = 1 | x, s = 1) &= 1 \\ p(y = x | x, s = 2) &= 1, \end{aligned} \tag{17}$$

where $\Pr\{S = 0\} = p$ is the probability that the channel is stuck at 0, $\Pr\{S = 1\} = q$ is the probability that the channel is stuck at 1, and $\Pr\{S = 2\} = r$ is the probability of having a good channel where $y = x$ with $p + q + r = 1$. We consider a binary symmetric channel (BSC) from Y to Z, where

$$Z = X \oplus N, \tag{18}$$

with $N \sim Bern(n)$ for some $n \in [0, 0.5]$. This corresponds to a degraded DMC model. (See Fig. 2.)

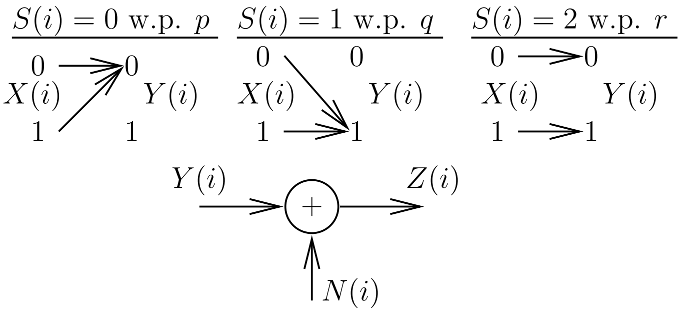

Fig. 2. Channel model of memory with defective cells. $p = \Pr\{S = 0\}$ (probability of being stuck at 0), $q = \Pr\{S = 1\} = q$ (probability of being stuck at 1), $r = \Pr\{S = 2\}$ (probability of being in a noiseless state), and $N \sim Bern(n)$, where $n \in [0, 0.5]$ is the cross over probability of the BSC from Y to Z.

We present numerical results for this channel model with three regions: Un-coded region, coded region, and an outer-bound region. The un-coded region is obtained by setting $U = \emptyset$ in Proposition 1, where we have the set of (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq I(S;Y) && (19)\\ R_l &\geq I(S;Z) && (20) \end{aligned}$$

over all possible $p(x|s)$. For the coded region, we simulate a sub-region of the one given in Proposition 3, where we set $U = Y$ and achieve the set of (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq \min\{H(S), H(Y)\} && (21)\\ R_l &\geq I(Y,S;Z) = H(Z) - H(N) && (22) \end{aligned}$$

over all possible $p(x|s)$. For converse arguments, we consider the outer-bound region given by the set of (R_a, R_l) pairs satisfying

$$\begin{aligned} R_a &\leq \min\{H(S), I(X,S;Y) = H(Y)\} && (23)\\ R_l &\geq I(S;Z) && (24) \end{aligned}$$

over all possible $p(x|s)$. This outer-bound region follows from Proposition 8. We evaluate the regions above in terms of the channel parameters as follows. Let $\Pr\{X = 1\} = \alpha$. Then,

$$\begin{aligned} H(S) &= H(p,q,r), && (25)\\ H(Y|S) &= rH(\alpha) && (26)\\ H(Y) &= H(q + r\alpha) && (27)\\ H(Z|S) &= (p+q)H(n) + rH(\alpha \otimes n) && (28)\\ H(Z) &= H((q + r\alpha) \otimes n), && (29) \end{aligned}$$

where $H(\cdot,\cdot,\cdot)$ is the ternary entropy function, $H(\cdot)$ is the binary entropy function, and $\otimes$ is the binary convolution given by $p \otimes q = p(1-q) + q(1-p)$. The numerical results are given in Fig. 3. The regions are truncated with $R_l \leq H(S)$ as any $R_l > H(S)$ is trivially achievable. We note that, the coded region is potentially larger than its un-coded counterparts even when we only compute a subset of the coded achievable region. This shows the enhancement provided by sending a refinement of the state sequence over the channel.

IV. Gaussian Scenario

Consider the channels given by

$$\begin{aligned} Y(i) &= X(i) + S(i) + N(i)\\ Z(i) &= X(i) + S(i) + N_z(i), && (30) \end{aligned}$$

where the state and noise distributions are generated i.i.d. as $S(i) \sim \mathcal{N}(0, \sigma_s^2)$, $N(i) \sim \mathcal{N}(0, \sigma_n^2)$, $N_z(i) \sim \mathcal{N}(0, \sigma_{n_z}^2)$, and the cost constraint on the channel input is given by $c(x) = x^2$ and $C = P$, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\{|X(i)|^2\} \leq P. \qquad (31)$$

(See Fig. 4.)

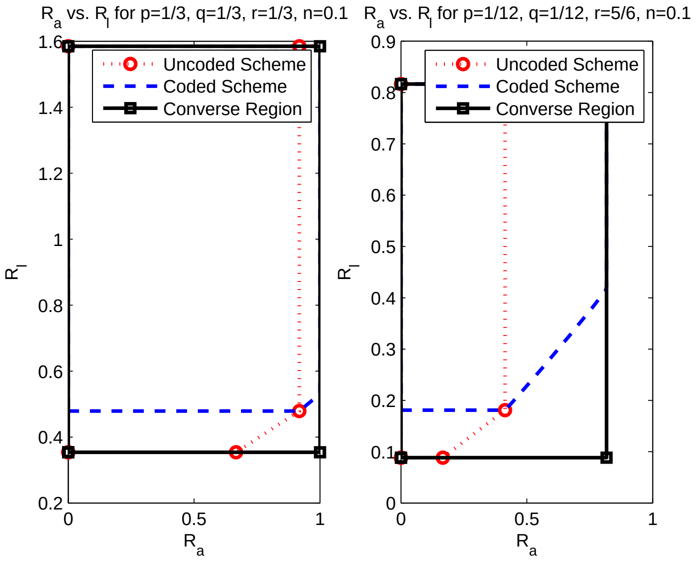

Fig. 3. Simulation results for memory with defective cells model.

A. *An inner bound for $\mathcal{C}$ using an un-coded scheme*

The inner bound is based on sending an amplified version of S together with some additional Gaussian noise. This *un-coded* signal is constructed as follows.

$$X(i) = \rho\frac{\sigma_x}{\sigma_s}S(i) + \sqrt{(1-\rho^2)}\sigma_x V(i), \tag{32}$$

where $V(i) \sim \mathcal{N}(0,1)$ independent of S(i), $\rho \in [-1,1]$, and $\sigma_x^2 \leq P$. Here, ρ^2 is the fraction of the power allocated to $S(i)$. This scheme achieves the following region.

Theorem 13: The un-coded scheme, which forwards $S(i)$ at each time step together with some i.i.d. Gaussian noise as given in (32), achieves the set of (R_a, R_l) pairs denoted by the region $\mathcal{R}^{\text{un-coded}}$, where

$$\mathcal{R}^{\text{un-coded}} = \text{Convex Hull}\left\{ \bigcup_{\rho\in[-1,1],\sigma_x^2\in[0,P]} (R_a(\rho,\sigma_x), R_l(\rho,\sigma_x)) \right\} \subset \mathcal{C},$$

with

$$\begin{aligned} R_a(\rho,\sigma_x) &= \frac{1}{2}\log\left(1+\frac{\sigma_s^2+2\rho\sigma_s\sigma_x+\rho^2\sigma_x^2}{\sigma_n^2+(1-\rho^2)\sigma_x^2}\right) \\ R_l(\rho,\sigma_x) &= \frac{1}{2}\log\left(1+\frac{\sigma_s^2+2\rho\sigma_s\sigma_x+\rho^2\sigma_x^2}{\sigma_{n_z}^2+(1-\rho^2)\sigma_x^2}\right). \end{aligned}$$

The expressions above are obtained by evaluating $R_a = I(S;Y)$ and $R_l = I(S;Z)$ on account of uncoded transmission in (32).

Examples:

- If $P \geq \sigma_s^2$, one can set $X = -S$ and achieve the pair

$$(R_a = 0, R_l = 0).$$

- Another trivial point is obtained by setting $X = 0$, which achieves

$$\left(R_a = \frac{1}{2}\log\left(1+\frac{\sigma_s^2}{\sigma_n^2}\right), R_l = \frac{1}{2}\log\left(1+\frac{\sigma_s^2}{\sigma_{n_z}^2}\right)\right).$$

B. *Outer-bounds on $\mathcal{C}$*

Corollary 14: Let ρ denote the correlation coefficient between X and S. The set of all achievable rate pairs (R_a, R_l), satisfy

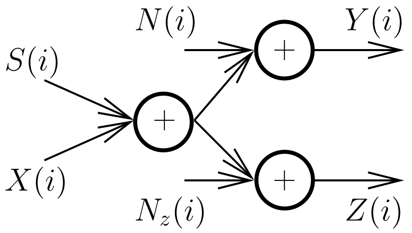

Fig. 4. The channel model for the Gaussian setting. $S(i) \sim \mathcal{N}(0, \sigma_s^2)$, $N(i) \sim \mathcal{N}(0, \sigma_n^2)$, and $N_z(i) \sim \mathcal{N}(0, \sigma_{n_z}^2)$.

$$\begin{aligned} R_a &\leq \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_n^2}\right) \\ R_l &\geq \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \rho^2\sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}\right) \end{aligned}$$

for $-1 \leq \rho \leq 1$ and $\sigma_x^2 \leq P$.

Proof: Using Proposition 9, we have

$$R_a \leq I(X,S;Y) = h(Y) - h(Y|X,S) = h(Y) - h(N) \leq \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2}{\sigma_n^2}\right). \tag{33}$$

Using Proposition 9, the linear estimate $\hat{S}(Z) = \frac{\mathbb{E}[SZ]}{\mathbb{E}[Z^2]}Z$ and the fact the conditioning reduces entropy, we get

$$R_l \geq I(S;Z,U) \geq I(S;Z) = h(S) - h(S|Z) = h(S) - h(S - \hat{S}(Z)|Z) \geq h(S) - h(S - \hat{S}(Z)). \tag{34}$$

Since the entropy maximizing distribution for a given second moment is a Gaussian, we have

$$h(S - \hat{S}(Z)) \leq \frac{1}{2}\log 2\pi e\left(\frac{\sigma_s^2}{1 + \frac{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \rho^2\sigma_x^2}{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}}\right), \tag{35}$$

leading to

$$R_l \geq \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \rho^2\sigma_x^2}{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}\right). \tag{36}$$

■

Corollary 15: Let ρ denote the correlation coefficient between X and S. If $\sigma_n^2 \leq \sigma_{n_z}^2$, then the set of all achievable rate pairs (R_a, R_l) satisfy

$$R_a - R_l \leq \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2}{\sigma_n^2}\right) - \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2}{\sigma_{n_z}^2}\right),$$

for $-1 \leq \rho \leq 1$ and $\sigma_x^2 \leq P$.

Proof: By Proposition 9, we have

$$R_a - R_l \leq I(X,S;Y|Z). \tag{37}$$

Without loss of generality, we consider $N_z = N + N_z'$ with $\sigma_{n_z}^2 = \sigma_n^2 + \sigma_{n'}^2$ where N_z' is independent of N. Noting that,

$$I(X,S;Y|Z) = h(Y|Z) - h(Y|X,S,Z) = h(Y|Z) - h(N|N_z), \tag{38}$$

we upper bound $h(Y|Z)$ using the following. Consider two zero-mean correlated random variables A and B.

$$\begin{aligned} h(A|B) &\overset{(a)}{=} h(A - \hat{A}(B)|B) \\ &\leq h(A - \hat{A}(B)) \\ &\overset{(b)}{\leq} \frac{1}{2}\log(2\pi e\sigma_e^2), \end{aligned} \tag{39}$$

where in (a) we used $\hat{A}(B)$ as the estimate of A given B, and (b) follows by defining the estimation error variance $\sigma_e^2 \triangleq E\left[(A - \hat{A}(B))^2\right]$ and the fact that Gaussian distribution maximizes entropy given the variance. We then upper bound the optimal estimator error variance by the linear MMSE variance. Therefore,

$$h(A|B) \leq \frac{1}{2}\log\left(2\pi e\left(\text{var}(A) - \frac{E\left[(AB)^2\right]}{\text{var}(B)}\right)\right). \tag{40}$$

Using the above, we obtain

$$\begin{aligned} R_a - R_l &\leq \frac{1}{2}\log\left(2\pi e\left(\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2 + \sigma_n^2 - \frac{(\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2 + \sigma_n^2)^2}{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2 + \sigma_n^2 + \sigma_{n'}^2}\right)\right) \\ &\quad - \frac{1}{2}\log\left(2\pi e\left(\sigma_n^2 - \frac{(\sigma_n^2)^2}{\sigma_n^2 + \sigma_{n'}^2}\right)\right) \\ &= \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2}{\sigma_n^2}\right) - \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + 2\rho\sigma_s\sigma_x + \sigma_x^2}{\sigma_n^2 + \sigma_{n'}^2}\right). \end{aligned} \tag{41}$$

This completes the proof. ■

C. Comparison of inner and outer bounds for the degraded Gaussian channel

We now compare the uncoded scheme and the outer bound presented above. In particular, we show that the uncoded transmission scheme achieves certain corner points of the amplification-masking region and that the gap between the inner and outer bounds on the region is within $1/2$ bit for certain channel parameters. We also show that the uncoded scheme achieves the optimal difference $R_a - R_l$.

1) Characterization of the gap between achievable and converse regions: We show that given any point (R_a, R_l) in the converse region corresponding to a given (ρ, σ_x), uncoded transmission achieves within $1/2$ bit of the converse region under certain conditions on channel parameters. In particular, for any given R_a, uncoded transmission achieves that R_a and within $1/2$ bit of the bound on R_l if $\frac{P}{\sigma_{n_z}^2} \leq 1$. Similarly for any given R_l, uncoded transmission achieves the given R_l and within $1/2$ bit of the bound on R_a if $\frac{P}{\sigma_n^2} \leq 1$. We prove these as follows. Using Corollary (14), any point in the outer bound region is described as

$$R_a = \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_n^2}\right) \tag{42}$$

$$R_l = \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \rho^2\sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}\right) \tag{43}$$

for $-1 \leq \rho \leq 1$ and $\sigma_x^2 \leq P$. Now let us show that uncoded transmission achieves any R_l in the region above and the gap from R_a as above is within $1/2$ bit. Let the uncoded scheme be designed such that $X_i = \frac{\sigma_x}{\sigma_s}\rho S_i + V_i$, where $V_i \sim \mathcal{N}(0, \sigma_x^2(1-\rho^2))$ and independent of S_i. Now, by (33) and (33) this input achieves a leakage, $I(S;Z) = \frac{1}{2}\log\left(1 + \frac{\sigma_s^2+\rho^2\sigma_x^2+2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2+\sigma_x^2(1-\rho^2)}\right)$ and R_a given by $I(S;Y) = \frac{1}{2}\log\left(1 + \frac{\sigma_s^2+\rho^2\sigma_x^2+2\rho\sigma_s\sigma_x}{\sigma_n^2+\sigma_x^2(1-\rho^2)}\right)$, which implies the gap is given by

$$I(X,S;Y) - I(S;Y) = I(X;Y|S) = \frac{1}{2}\log\left(1 + \frac{\sigma_x^2(1-\rho^2)}{\sigma_n^2}\right) \leq \frac{1}{2}, \tag{44}$$

for $\frac{P}{\sigma_n^2} < 1$. Now, in order to prove the other claim, that uncoded achieves any given R_a and the gap with R_l is within $1/2$ bit, we proceed as follows. Given $R_a = \frac{1}{2}\log\left(1 + \frac{\sigma_s^2+\sigma_x^2+2\rho\sigma_s\sigma_x}{\sigma_n^2}\right)$, we achieve this by choosing an uncoded scheme such that $X_i = \frac{\sigma_{x'}}{\sigma_s}S_i$ if $\rho > 0$ or $X_i = -\frac{\sigma_{x'}}{\sigma_s}S_i$ if $\rho \leq 0$, and

$$\sigma_{x'}^2 + 2\frac{\rho}{|\rho|}\sigma_s\sigma_{x'} = \sigma_x^2 + 2\rho\sigma_s\sigma_x, \tag{45}$$

where $0 \leq \sigma_{x'} \leq \sqrt{P}$. By the intermediate value theorem for continuous functions, it is clear that there exists $\sigma_{x'} \leq \sqrt{P}$ such that the condition above is satisfied. Further, the uncoded scheme achieves the desired R_a. The scheme achieves an R_l given by $\frac{1}{2}\log\left(1 + \frac{\sigma_s^2+\sigma_{x'}^2+2\frac{\rho}{|\rho|}\sigma_s\sigma_{x'}}{\sigma_{n_z}^2}\right)$ leading to a gap

$$\begin{aligned} &\frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \sigma_{x'}^2 + 2\frac{\rho}{|\rho|}\sigma_s\sigma_{x'}}{\sigma_{n_z}^2}\right) - \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \rho^2\sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}\right) \\ &= \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2}\right) - \frac{1}{2}\log\left(1 + \frac{\sigma_s^2 + \rho^2\sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}\right) \\ &= \frac{1}{2}\log\left(\frac{\sigma_{n_z}^2 + \sigma_s^2 + \sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2}\right) - \frac{1}{2}\log\left(\frac{\sigma_{n_z}^2 + \sigma_s^2 + \sigma_x^2 + 2\rho\sigma_s\sigma_x}{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}\right) \\ &\leq \frac{1}{2}\log\left(\frac{\sigma_{n_z}^2 + \sigma_x^2(1-\rho^2)}{\sigma_{n_z}^2}\right) = \frac{1}{2}\log\left(1 + \frac{\sigma_x^2(1-\rho^2)}{\sigma_{n_z}^2}\right) \leq \frac{1}{2}\log 2 = \frac{1}{2}, \end{aligned} \tag{46}$$

when $\frac{P}{\sigma_{n_z}^2} < 1$.

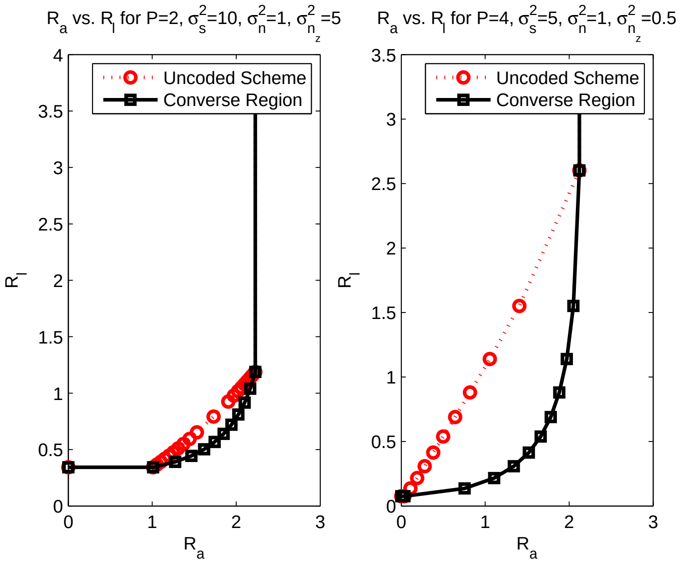

Fig. 5. Simulation results for the Gaussian scenario.

2) Differential amplification capacity: Note that the un-coded transmission achieves the maximum $R_a - R_l$. The upper bound on $R_a - R_l$ in (15) is maximized for $\sigma_x^2 = P$ and $\rho = 1$. Thus, this maximum difference between R_a and R_l is achieved by un-coded transmission corresponding to $X = \frac{\sqrt{P}}{\sigma_s} S$ in Theorem 13, and is given by

$$C_d = \frac{1}{2}\log\left(1 + \frac{(\sigma_s + \sqrt{P})^2}{\sigma_n^2}\right) - \frac{1}{2}\log\left(1 + \frac{(\sigma_s + \sqrt{P})^2}{\sigma_{n_z}^2}\right) \tag{47}$$

3) Corner points of the trade-off region: Consider the corner points of the amplification-masking region. Inspecting (33), we observe that the point in the outer bound region corresponding to maximum amplification is given by $\rho = 1$. Clearly, from (44) and (46), we see that the gap is zero for $\rho = 1$. Similarly, consider the point corresponding to minimum leakage R_l in the *weak* and *moderate* interference regimes as in [30]. These points again correspond to $\rho = -1$ and we have $I(X, S; Y) = I(S; Y)$ and $I(X, S; Z) = I(S; Z)$, leading to the gap being zero. This is also verified by setting $\rho = -1$ in (44) and (46).

4) Numerical results: We compare the un-coded region with an outer-bound region (in Proposition 14) in Fig. 5. The first case corresponds to a degraded scenario, where the gap between the regions is fairly small as expected from the analysis given above. However, for the reversely degraded scenario with larger power constraint P compared to the state power σ_s^2, the gap is larger. In Fig. 6, we plot the differential amplification capacity for a degraded channel ($\sigma_n^2 = 1$, $\sigma_{n_z}^2 = 5$) for a range of power constraints P and different values of σ_s^2. Note that the differential amplification capacity saturates in the high SNR regime, and the effect of encoder in increasing C_d is decreasing as the power of the additive state increases.

V. Conclusion

We study the problem of state amplification under the masking constraints, where the encoder (with the knowledge of non-causal state S^n) facilitates the amplification rate ($\frac{1}{n}I(S^n; Y^n)$) at Bob (observing Y^n) while minimizing the leakage rate ($\frac{1}{n}I(S^n; Z^n)$) as much as possible at Eve (observing Z^n). Our coding schemes are based on indexing the state sequence and sending one of the indices over the channel to Bob. The achievable region corresponding to this strategy is derived by calculating bounds on amplification and masking rates. We also show that for the input distributions enabling Bob to be a “stronger” receiver than Eve, the index of the state can be sent securely over the channel. This **secure refinement** approach is shown to lead to non-trivial achievable regions. We also provided outer bounds, using which we showed that the scheme without secure refinement achieves the optimal $R_a - R_l$ over the region in the reversely degraded DMCs, the degraded binary channels, and Gaussian channels. For the degraded Gaussian model, we also characterized the optimal corner points, and the gap between the outer bound and achievable regions.

Several interesting problems can be considered as future directions. First, the channel cost may be introduced for the DMC model as well, and the cost may have some dependence on the state sequence or vary according to a stochastic model. Second, causal channel state knowledge can be considered. In addition, one can also consider sending messages to receivers in addition to the task of state amplification and masking. Our current focus is on such extensions. Also, the coded state sequence setting [34] (a scenario that is more relevant to the cognitive radio type systems, where the primary signal carrying a message corresponds to the channel state sequence that is non-causally known at the cognitive encoder) may also be considered.

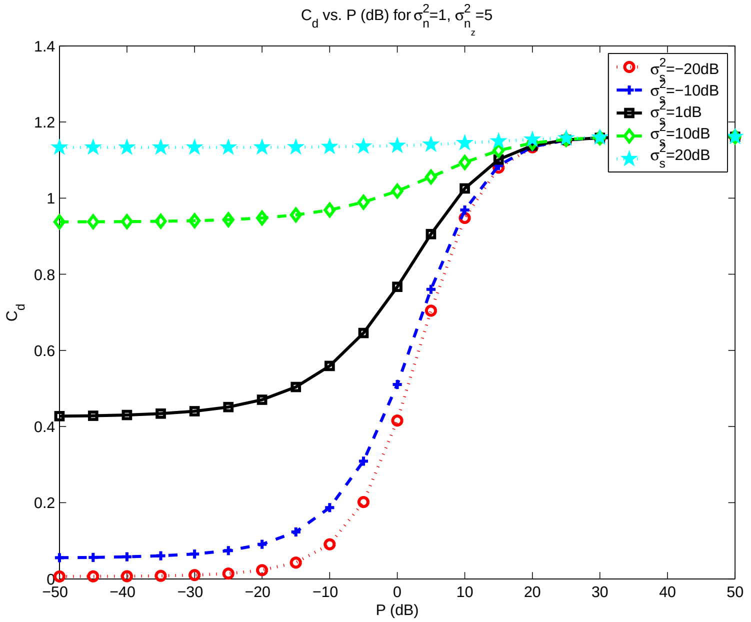

Fig. 6. Differential amplification capacity C_d vs. power P (dB). The dashed curve with diamond markers correspond to the same scenario given in the degraded setting of Fig. 5 ($\sigma_s^2 = 10$) for different power levels.

Appendix A
Proof of Proposition 1

Proof: Fix $p(u, x|s)$, and consider s^n, the state sequence of the channel that is non-causally known at the encoder. We generate 2^{nR_u} codewords denoted by $u^n(w_u)$ with $w_u \in \{1, \cdots, 2^{nR_u}\}$, each distributed according to $\prod_{i=1}^{n} p(u_i)$.

The encoder chooses a $u^n(k)$ such that $(u^n(k), s^n) \in \mathcal{T}_\epsilon^n$. ($\epsilon$-typical sets are denoted by $\mathcal{T}_\epsilon^n$, and standard definition and properties of typical sets are used throughout the text [36].) If no such codeword exists, an arbitrary sequence is picked. The encoder sends x^n generated by $\prod_{i=1}^{n} p(x_i|u_i(k), s_i)$.

We now derive an achievable amplification rate with this scheme. We consider the following.

$$\frac{1}{n}I(S^n;Y^n) \overset{(a)}{=} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1)(1-\Pr\{\mathcal{E}_1^c\}) + \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1^c)\Pr\{\mathcal{E}_1^c\} \tag{48}$$

$$\geq \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1)(1-\Pr\{\mathcal{E}_1^c\}) \tag{49}$$

$$\overset{(b)}{\geq} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1) - \Pr\{\mathcal{E}_1^c\}(H(S)+\epsilon_1) \tag{50}$$

$$\overset{(c)}{=} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1,\mathcal{E}_2)(1-\Pr\{\mathcal{E}_2^c|\mathcal{E}_1\}) + \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1,\mathcal{E}_2^c)\Pr\{\mathcal{E}_2^c|\mathcal{E}_1\} - \Pr\{\mathcal{E}_1^c\}(H(S)+\epsilon_1) \tag{51}$$

$$\overset{(d)}{\geq} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1,\mathcal{E}_2) - (\Pr\{\mathcal{E}_1^c\} + \Pr\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_1) \tag{52}$$

$$= \frac{1}{n}I(S^n;Y^n,U^n|\mathcal{E}_1,\mathcal{E}_2) - (\Pr\{\mathcal{E}_1^c\} + \Pr\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_1) \tag{53}$$

$$\overset{(e)}{\geq} (H(S)-\epsilon_1) - (H(S|Y,U)+\epsilon_2) - (\Pr\{\mathcal{E}_1^c\} + \Pr\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_1) \tag{54}$$

$$\overset{(f)}{=} I(S;Y,U) - \hat{\epsilon}_1 \tag{55}$$

where in (a) $\mathcal{E}_1$ is the event that S^n is a typical sequence, (b) follows as $\frac{1}{n}I(S^n;Y^n|\mathcal{E}_1) \leq \frac{1}{n}H(S^n|\mathcal{E}_1) \leq H(S)+\epsilon_1$ as the number of typical S^n sequences are bounded above by $2^{n(H(S)+\epsilon_1)}$ with $\epsilon_1 \to 0$ as $n \to \infty$, in (c) $\mathcal{E}_2$ is the event that U^n is decoded given Y^n, (d) is similar to (b), (e) follows as $H(S^n|\mathcal{E}_1)$ is lower bounded by $n(H(S)-\epsilon_1)$ and $H(S^n|Y^n,U^n,\mathcal{E}_1,\mathcal{E}_2)$ is upper bounded by $n(H(S|Y,U)+\epsilon_2)$ as U^n, S^n, Y^n are jointly typical, with $\epsilon_2 \to 0$ as $n \to \infty$, in (f) we define $\hat{\epsilon}_1 \triangleq (\Pr\{\mathcal{E}_1^c\} + \Pr\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_1) + \epsilon_1 + \epsilon_2$.

Here, as $n \to \infty$, $\Pr\{\mathcal{E}_1^c\} \to 0$, and $\Pr\{\mathcal{E}_2^c|\mathcal{E}_1\} \to 0$ when we select

$$\begin{aligned} R_u &= I(U;S) + \epsilon_3 && (56)\\ R_u &< I(U;Y), && (57)\end{aligned}$$

where the first condition makes encoding error arbitrarily small (also known as the mutual covering lemma), and the second one allows obtaining U^n jointly typical with Y^n given that there is no encoding error, i.e., U^n, S^n are jointly typical and generates Y^n. This shows that $\frac{1}{n}I(S^n;Y^n) \geq I(S;Y,U) - \hat{\epsilon}_1 \geq R_a - \hat{\epsilon}_1$, i.e., any $R_a \leq I(S;Y,U)$ is achievable.

We now derive the achievable leakage rate expression for this scheme.

$$\begin{aligned} \frac{1}{n}I(S^n;Z^n) &\overset{(a)}{=} \frac{1}{n}I(S^n;Z^n|\mathcal{E}_1)(1-\Pr\{\mathcal{E}_1^c\}) + \frac{1}{n}I(S^n;Z^n|\mathcal{E}_1^c)\Pr\{\mathcal{E}_1^c\} && (58)\\ &\leq \frac{1}{n}I(S^n;Z^n|\mathcal{E}_1) + H(S)\Pr\{\mathcal{E}_1^c\} && (59)\\ &\overset{(b)}{\leq} \frac{1}{n}I(S^n;Z^n,U^n|\mathcal{E}_1) + H(S)\Pr\{\mathcal{E}_1^c\} && (60)\\ &\overset{(c)}{\leq} (H(S)+\epsilon_1) - (H(S|Z,U)-\epsilon_2) + H(S)\Pr\{\mathcal{E}_1^c\} && (61)\\ &\overset{(d)}{=} I(S;Z,U) + \hat{\epsilon}_2 && (62)\end{aligned}$$

where in (a) $\mathcal{E}_1$ is the event that S^n is a typical sequence, in (b) we included the U^n chosen at the encoder, (c) follows as $H(S^n|\mathcal{E}_1)$ is upper bounded by $n(H(S)+\epsilon_1)$ and $H(S^n|Z^n,U^n,\mathcal{E}_1)$ is lower bounded by $n(H(S|Z,U)-\epsilon_2)$, which can be shown by using the event that U^n, S^n, Z^n are jointly typical w.h.p., and in (d) we define $\hat{\epsilon}_2 \triangleq \epsilon_1 + \epsilon_2 + H(S)\Pr\{\mathcal{E}_1^c\}$. Noting that $\hat{\epsilon}_2 \to 0$ as $n \to \infty$ concludes the proof as $\frac{1}{n}I(S^n;Z^n) \leq I(S;Z,U) + \hat{\epsilon}_2 \leq R_l + \hat{\epsilon}_2$, i.e., any $R_l \geq I(S;Z,U)$ is achievable. Along the same lines, one can similarly obtain that any $R_l \geq I(U,S;Z)$ is achievable. (Also given in the proof of Proposition 3.) ■

Appendix B
Proof of Proposition 3

Proof: Fix $p(u,x|s)$, and consider s^n, the state sequence of the channel that is non-causally known at the encoder. We generate $2^{n(R_u+R'_u)}$ codewords denoted by $U^n(w_u,w'_u)$ each distributed according to $\prod_{i=1}^{n} p(u_i)$, where $w_u \in \{1,\cdots,2^{nR_u}\}$ and $w'_u \in \{1,\cdots,2^{nR'_u}\}$. We also list all the typical S^n sequences with two indices $S^n(w_u,w_r)$.

For the given s^n sequence, if it is a typical sequence the encoder identify it as $s^n(k,l)$, otherwise arbitrary indices are chosen. Encoder then finds the index $m \in \{1,\cdots,2^{nR'_u}\}$, such that $u^n(k,m)$ and $s^n(k,l)$ are jointly typical. If no such codeword exists, an arbitrary sequence is picked. The encoder sends x^n generated by $\prod_{i=1}^{n} p(x_i|u_i(k,m),s_i(k,l))$.

Denote $\mathcal{E}_1$ to be the event that S^n is a typical sequence, and $\mathcal{E}_2$ is the event that U^n is decoded given Y^n using joint typicality decoder. Here, as $n \to \infty$, $\Pr\{\mathcal{E}_1^c\} \to 0$, and $\Pr\{\mathcal{E}_2^c|\mathcal{E}_1\} \to 0$ when we select

$$\begin{aligned} R'_u &= I(U;S) + \epsilon_1 && (63)\\ R_u + R'_u &\leq I(U;Y), && (64)\end{aligned}$$

where the first condition makes encoding error arbitrarily small (see [18]), and the second one allows decoding $U^n(K,M)$ using joint typicality with Y^n given that there is no encoding error, i.e., $U^n(K,M), S^n(K,L)$ are jointly typical and generates Y^n. We set

$$R_u = \min\{I(U;Y) - I(U;S) - \epsilon_1, H(S|U,Y) - \epsilon_1\} \quad (65)$$

with some $\epsilon_1 \to 0$ as $n \to \infty$.

We derive the achievable amplification rate as follows.

$$\begin{aligned} \frac{1}{n}I(S^n;Y^n) &\overset{(a)}{=} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1)(1-\Pr\{\mathcal{E}_1^c\}) + \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1^c)\Pr\{\mathcal{E}_1^c\} && (66)\\ &\geq \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1)(1-\Pr\{\mathcal{E}_1^c\}) && (67)\\ &\overset{(b)}{\geq} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1) - \Pr\{\mathcal{E}_1^c\}(H(S)+\epsilon_2) && (68)\\ &\overset{(c)}{=} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1,\mathcal{E}_2)(1-\Pr\{\mathcal{E}_2^c|\mathcal{E}_1\}) + \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1,\mathcal{E}_2^c)\Pr\{\mathcal{E}_2^c|\mathcal{E}_1\}\\ &\quad - \Pr\{\mathcal{E}_1^c\}(H(S)+\epsilon_2) && (69)\end{aligned}$$

$$\overset{(d)}{\geq} \frac{1}{n}I(S^n;Y^n|\mathcal{E}_1,\mathcal{E}_2)-(\text{Pr}\{\mathcal{E}_1^c\}+\text{Pr}\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_2) \tag{70}$$

$$= \frac{1}{n}I(S^n(K,L);Y^n,U^n(K,M)|\mathcal{E}_1,\mathcal{E}_2) - (\text{Pr}\{\mathcal{E}_1^c\}+\text{Pr}\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_2) \tag{71}$$

$$\begin{aligned} = {} & \frac{1}{n}H(S^n(K,L)|\mathcal{E}_1,\mathcal{E}_2)-\frac{1}{n}H(S^n(K,L)|U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) \\ & +\frac{1}{n}H(Y^n|U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) \\ & -\frac{1}{n}H(Y^n|S^n(K,L),U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) \\ & -(\text{Pr}\{\mathcal{E}_1^c\}+\text{Pr}\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_2) \end{aligned} \tag{72}$$

$$\begin{aligned} \overset{(e)}{\geq} {} & R_u+R_r-R_r+I(U;S)-\epsilon_1+(H(Y|U)-\epsilon_2)-(H(Y|U,S)+\epsilon_3) \\ & -(\text{Pr}\{\mathcal{E}_1^c\}+\text{Pr}\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_2) \end{aligned} \tag{73}$$

$$\overset{(f)}{=} \min\{H(S),I(U,S;Y)\}-\hat{\epsilon}_1 \tag{74}$$

where in (a) we used the event $\mathcal{E}_1$, (b) follows as $\frac{1}{n}I(S^n;Y^n|\mathcal{E}_1)\leq\frac{1}{n}H(S^n|\mathcal{E}_1)\leq H(S)+\epsilon_2$ as the number of typical S^n sequences are bounded above by $2^{n(H(S)+\epsilon_2)}$ with $\epsilon_2\to 0$ as $n\to\infty$, in (c) we used the event $\mathcal{E}_2$, (d) is similar to (b), (e) follows as

$$\begin{aligned} H(S^n(K,L)|\mathcal{E}_1,\mathcal{E}_2) &= n(R_u+R_r) && (75)\\ H(S^n(K,L)|U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) &\leq n(R_r-I(U;S)+\epsilon_1) && (76)\\ H(Y^n|U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) &\geq n(H(Y|U)-\epsilon_2) && (77)\\ H(Y^n|S^n(K,L),U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) &\leq n(H(Y|S,U)+\epsilon_3), && (78) \end{aligned}$$

where $\epsilon_2,\epsilon_3\to 0$ as $n\to\infty$. (We show (76) in Lemma 16 and (77) in Lemma 17, both are given at the end of this proof. The last bound can be shown by using the event that $(U^n(K,M),S^n(K,L),Y^n)$ is jointly typical w.h.p. as each Y_i is generated from $s_i(k,l),u_i(k,m)$ with the associated $p(y|s,u)$.) In (f), we define $\hat{\epsilon}_1\triangleq(\text{Pr}\{\mathcal{E}_1^c\}+\text{Pr}\{\mathcal{E}_2^c|\mathcal{E}_1\})(H(S)+\epsilon_2)+\epsilon_1+\epsilon_2+\epsilon_3$, where readily have $\hat{\epsilon}_1\to 0$ as $n\to\infty$.

This shows that $\frac{1}{n}I(S^n;Y^n)\geq\min\{H(S),I(U,S;Y)\}-\hat{\epsilon}_1\geq R_a-\hat{\epsilon}_1$, i.e., any

$$R_a\leq\min\{H(S),I(U,S;Y)\} \tag{79}$$

is achievable.

We bound $\frac{1}{n}I(S^n;Z^n)$ as follows.

$$\begin{aligned} \frac{1}{n}I(S^n;Z^n) &\overset{(a)}{=} \frac{1}{n}I(S^n;Z^n|\mathcal{E}_1)(1-\text{Pr}\{\mathcal{E}_1^c\})+\frac{1}{n}I(S^n;Z^n|\mathcal{E}_1^c)\text{Pr}\{\mathcal{E}_1^c\} && (80)\\ &\leq \frac{1}{n}I(S^n;Z^n|\mathcal{E}_1)+H(S)\text{Pr}\{\mathcal{E}_1^c\} && (81)\\ &\overset{(b)}{\leq} \frac{1}{n}I(U^n(K,M),S^n(K,L);Z^n|\mathcal{E}_1)+H(S)\text{Pr}\{\mathcal{E}_1^c\} && (82)\\ &\overset{(c)}{\leq} (H(Z)+\epsilon_1)-(H(Z|U,S)-\epsilon_2)+H(S)\text{Pr}\{\mathcal{E}_1^c\} && (83)\\ &\overset{(d)}{=} I(U,S;Z)+\hat{\epsilon}_2 && (84) \end{aligned}$$

where in (a) $\mathcal{E}_1$ is the event that S^n is a typical sequence, in (b) we included the $U^n(K,M)$ chosen at the encoder which observes the typical sequence $S^n(K,L)$ (c) follows as $H(Z^n|\mathcal{E}_1)$ is upper bounded by $n(H(Z)+\epsilon_1)$ and the term $H(Z^n|U^n(K,M),S^n(K,L),\mathcal{E}_1)$ is lower bounded by $n(H(Z|U,S)-\epsilon_2)$, which can be shown by using the event that $U^n(K,M),S^n(K,L),Z^n$ are jointly typical w.h.p. for any $K=w_u$. (Note that this implies that Z^n is a typical sequence w.h.p. as $(U^n(K,M),S^n(K,L))$ is typical.) In (d), we define $\hat{\epsilon}_2\triangleq\epsilon_1+\epsilon_2+H(S)\text{Pr}\{\mathcal{E}_1^c\}$. Noting that $\hat{\epsilon}_2\to 0$ as $n\to\infty$ concludes the proof as $\frac{1}{n}I(S^n;Z^n)\leq I(U,S;Z)+\hat{\epsilon}_2\leq R_l+\hat{\epsilon}_2$, i.e., any $R_l\geq I(U,S;Z)$ is achievable.

Lemma 16: For (76), we can obtain the bound using the following argument.

$$\begin{aligned}
H(S^n(K,L)|U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) &= H(S^n(K,L)|K,U^n(K,M),\mathcal{E}_1,\mathcal{E}_2) && (85)\\
&= H(S^n(K,L)|K,\mathcal{E}_1,\mathcal{E}_2) - I(S^n(K,L);U^n(K,M)|K,\mathcal{E}_1,\mathcal{E}_2) && (86)\\
&= nR_r - \sum_{w_u\in\{1,\cdots,2^{nR_u}\}} \Pr\{K=w_u\} I(S^n(K,L);U^n(K,M)|K=w_u,\mathcal{E}_1,\mathcal{E}_2) && (87)\\
&= nR_r - \sum_{w_u\in\{1,\cdots,2^{nR_u}\}} \Pr\{K=w_u\}\Bigg(H(U^n(K,M)|K=w_u,\mathcal{E}_1,\mathcal{E}_2)\\
&\qquad - H(U^n(K,M)|K=w_u,S^n(K,L),\mathcal{E}_1,\mathcal{E}_2)\Bigg) && (88)\\
&\overset{(a)}{\leq} nR_r - \sum_{w_u\in\{1,\cdots,2^{nR_u}\}} \Pr\{K=w_u\} n\left(I(U;S)-\epsilon_1\right) && (89)\\
&= n(R_r - I(U;S) + \epsilon_1) && (90)
\end{aligned}$$

where (a) holds as $H(U^n(K,M)|K=w_u,S^n(K,L),\mathcal{E}_1,\mathcal{E}_2)=0$, as M is determined when there is no encoding error, and $H(U^n(K,M)|K=w_u,\mathcal{E}_1,\mathcal{E}_2)\geq n(I(U;S)-\epsilon_1)$ with $\epsilon_1\to 0$ as $n\to\infty$, which follows as each $U^n(w_u,m)$ is a typical sequence given $\mathcal{E}_1$ and $\mathcal{E}_2$, where each codeword is approximately uniformly distributed; and the number of $U^n(w_u,M)$ codewords is $2^{n(I(U;S)+\epsilon_1)}$.

Lemma 17: Note that the event $\mathcal{E}_2$ implies that Y^n is jointly typical with the decoded sequence $U^n(K,M)$. Therefore, it follows that

$$H(Y^n|U^n(K,M),\mathcal{E}_1,\mathcal{E}_2)\geq n(H(Y|U)-\epsilon_2), \tag{91}$$

for some $\epsilon_2\to 0$ as $n\to\infty$.

■

Appendix C
Proof of Proposition 5

Proof: Consider a $p(u,x|s)$ satisfying $I(U;Y)\geq I(U;Z)$. We divide the proof into two cases.

(A) $I(U;Z)\geq I(U;S)$**:** We use the same codebook given in Proposition 3. The amplification rate analysis follows from Proposition 3.

We first consider the following bound for the leakage rate.

$$\begin{aligned}
H(W_u|Z^n) &= H(W_u,W_u',Z^n) - H(W_u'|W_u,Z^n) - H(Z^n) && (92)\\
&= H(U^n,Z^n) + H(W_u,W_u'|U^n,Z^n) - H(W_u'|W_u,Z^n) - H(Z^n) && (93)\\
&\overset{(a)}{\geq} n(R_u + I(U;S) - I(U;Z) - \epsilon_4), && (94)
\end{aligned}$$

where (a) follows as $H(W_u'|W_u,Z^n)\leq n\epsilon_2$ due to the fact that, given W_u, Z^n can decode W_u' from the codebook $U^n(W_u,W_u')$, $H(W_u,W_u'|U^n,Z^n)\geq 0$, and $H(U^n|Z^n)=H(U^n)-I(U^n;Z^n)=n(R_u+R_u')-I(U^n;Z^n)\geq n(R_u+R_u'-I(U;Z)-\epsilon_3)=n(R_u+I(U;S)-I(U;Z)-\epsilon_3+\epsilon_1)$ due to $R_u'=I(U;S)+\epsilon_1$ and bounding $I(U^n;Z^n)\leq n(I(U;Z)+\epsilon_3)$. We combined ϵ_k terms within some $\epsilon_4\to 0$ as $n\to\infty$.

Using this, we can get the following

$$I(W_u;Z^n)=H(W_u)-H(W_u|Z^n)\leq n(\min\{R_u,I(U;Z)-I(U;S)\}+\epsilon_4), \tag{95}$$

where the first bound in minimum expression is the trivial upper bound.

Then, we bound the leakage rate using the event $\mathcal{E}_1$: The event that S^n is typical. (We suppress all diminishing terms within some $\epsilon_k\to 0$ as $n\to\infty$.)

$$\begin{aligned}
I(S^n;Z^n) &\leq I(S^n(W_u,W_r);Z^n|\mathcal{E}_1)+n\epsilon_5 && (96)\\
&= I(W_u,S^n(W_u,W_r);Z^n|\mathcal{E}_1)+n\epsilon_5 && (97)\\
&= I(W_u;Z^n|\mathcal{E}_1)+H(S^n(W_u,W_r)|W_u,\mathcal{E}_1)-H(S^n(W_u,W_r)|Z^n,W_u,\mathcal{E}_1)+n\epsilon_5 && (98)\\
&\overset{(a)}{\leq} n\Bigg(\min\Big\{R_u,I(U;Z)-I(U;S)\Big\}+H(S)-R_u-H(S|Z,U)+R_u+\epsilon_6\Bigg) && (99)\\
&= n(I(S;Z,U)+\min\{R_u,I(U;Z)-I(U;S)\}+\epsilon_6) && (100)
\end{aligned}$$

where (a) follows from three observations:

1) We have $I(W_u;Z^n) = \text{Pr}\{\mathcal{E}_1\}I(W_u;Z^n|\mathcal{E}_1) + \text{Pr}\{\mathcal{E}_1^c\}I(W_u;Z^n|\mathcal{E}_1^c)$, which implies $I(W_u;Z^n|\mathcal{E}_1) \leq I(W_u;Z^n) + n\epsilon_7$, where the right hand side is upper bounded by (95),

2) $H(S^n(W_u,W_r)|W_u,\mathcal{E}_1) = n(H(S) - R_u)$, and

3) We observe

$$\begin{aligned} H(S^n(W_u,W_r)|Z^n,W_u,\mathcal{E}_1) &\geq H(S^n(W_u,W_r)|Z^n,U^n(W_u,W_u'),\mathcal{E}_1,\mathcal{E}_2) && (101)\\ &= H(S^n(W_u,W_r)|U^n(W_u,W_u'),\mathcal{E}_1,\mathcal{E}_2) - H(Z^n(W_u,W_r)|U^n(W_u,W_u'),\mathcal{E}_1,\mathcal{E}_2)\\ &\quad + H(Z^n|S^n(W_u,W_r),U^n(W_u,W_u'),\mathcal{E}_1,\mathcal{E}_2) && (102)\\ &\geq n(R_r - I(U;S) - H(Z|U) + H(Z|U,S) - \epsilon_8) && (103)\\ &= n(H(S|Z,U) - R_u - \epsilon_8), && (104)\end{aligned}$$

where we condition the entropy with the codeword chosen at the encoder together with the event $\mathcal{E}_2$ (defined as in the previous proof, but here with Z^n) in the first inequality, and the second inequality follows from the analysis given in the proof of Proposition 3: Using the same typicality arguments given for (77) and (78), we obtain

$$H(Z^n(W_u,W_r)|U^n(W_u,W_u'),\mathcal{E}_1,\mathcal{E}_2) \leq n(H(Z|U) + \epsilon_9) \tag{105}$$

and

$$H(Z^n|S^n(W_u,W_r),U^n(W_u,W_u'),\mathcal{E}_1,\mathcal{E}_2) \geq n(H(Z|S,U) - \epsilon_{10}), \tag{106}$$

and modifying the inequality in Lemma 16, we obtain

$$H(S^n(W_u,W_r)|U^n(W_u,W_u'),\mathcal{E}_1,\mathcal{E}_2) \geq n(R_r - I(U;S) - \epsilon_{11}) \tag{107}$$

similar to the single letterization of (76).

Therefore, the leakage rate

$$R_l \geq I(S;Z,U) + \min\{I(U;Z) - I(U;S), R_u\} \tag{108}$$

is achievable.

(B) $I(U;S) \geq I(U;Z)$**:** (Sketch of the proof) We modify the codebook given in Proposition 3. We design $U^n(W_u,W_u',W_u'')$ sequences where each index has the rate R_u, R_u', R_u'' respectively. We set $R_u' = I(U;S) - I(U;Z)$, $R_u'' = I(U;Z)$, and set the rate R_u as in the Proposition 3. Here, given w_u the encoder chooses w_u', w_u'' such that $u^n(w_u,w_u',w_u'')$ is jointly typical with $s^n(w_u,w_r)$.

Similar to the analysis above for obtaining (95), by replacing W_u with (W_u,W_u'), and W_u' with W_u'' in the analysis given in (93), we obtain $H(W_u,W_u'|Z^n) \geq n(R_u + R_{u'} - \epsilon_{12})$ and hence

$$I(W_u,W_u';Z^n) \leq n\epsilon_{12}, \tag{109}$$

showing that $I(W_u;Z^n) \leq n\epsilon_1$, and hence the message W_u is secured. Then, the leakage analysis given in part (A) above follows, where, using (109), we obtain

$$I(S^n;Z^n) \leq n(I(S;Z,U) + \epsilon_{13}). \tag{110}$$

The amplification rate analysis follows from Proposition 3. The leakage rate expression follows from the analysis given in the proof of Proposition 3 (showing the achievability of leakage $R_l \geq I(U,S;Z)$), in addition to the analysis above.

■

Appendix D
Proof of Proposition 8

Proof: We have the following bound.

$$\begin{aligned} I(S^n;Y^n) &\leq I(X^n,S^n;Y^n) && (111)\\ &\overset{(a)}{\leq} \sum_{i=1}^{n} I(X_i,S_i;Y_i) && (112)\\ &\overset{(b)}{=} nI(X,S;Y), && (113)\end{aligned}$$

where (a) is due to the memoryless channel $p(y_i|x_i,s_i)$ and the fact that conditioning does not increase the entropy, and (b) follows as each instant the channel is given by $p(y|x,s)$. In addition,

$$\begin{aligned} I(S^n;Y^n) &\leq H(S^n) && (114)\\ &= nH(S). && (115)\end{aligned}$$

We then bound $\frac{1}{n}I(S^n;Z^n)$.

$$\begin{aligned} I(S^n;Z^n) &= H(S^n)-H(S^n|Z^n) && (116)\\ &= \sum_{i=1}^{n} H(S_i)-H(S_i|S_1^{i-1},Z^n) && (117)\\ &\overset{(a)}{\geq} \sum_{i=1}^{n} H(S_i)-H(S_i|Z_i,S_1^{i-1},Z_{i+1}^n) && (118)\\ &\overset{(b)}{=} \sum_{i=1}^{n} H(S_i)-H(S_i|Z_i,U_i) && (119)\\ &\overset{(c)}{=} nI(S;Z,U), && (120)\end{aligned}$$

where (a) is due the fact that conditioning does not increase the entropy, in (b) we defined $U_i=(S_1^{i-1},Z_{i+1}^n)$.

We now obtain another condition using the definition of $U_i=(S_1^{i-1},Z_{i+1}^n)$.

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} I(Z_{i+1}^n;Z_i) && (121)\\ &= \sum_{i=1}^{n} I(S_1^{i-1},Z_{i+1}^n;Z_i)-I(S_1^{i-1};Z_i|Z_{i+1}^n) && (122)\\ &\overset{(a)}{=} \sum_{i=1}^{n} I(S_1^{i-1},Z_{i+1}^n;Z_i)-I(Z_{i+1}^n;S_i|S_1^{i-1}) && (123)\\ &= \sum_{i=1}^{n} I(S_1^{i-1},Z_{i+1}^n;Z_i)-I(S_1^{i-1},Z_{i+1}^n;S_i) && (124)\\ &= n(I(U;Z)-I(U;S)), && (125)\end{aligned}$$

where (a) follows by Csiszár's sum lemma. (We note the similarity of the analysis above with the converse result of the Gel'fand-Pinsker problem given in [37], where the indices of S and Z are reversed.)

We combine the bounds above with the fact that any achievable (R_a,R_l) for the given channel $p(y,z|x,s)$ and $p(s)$ should satisfy $\frac{1}{n}I(S^n;Y^n)\geq R_a-\epsilon$ and $\frac{1}{n}I(S^n;Z^n)\leq R_l+\epsilon$.

■

Appendix E
Proof of Proposition 9

Proof: Let $\mathcal{P}_1$ denote the set of $p(u,x|s)$ satisfying $I(U;Y)\geq I(U;S)$, and denote $\mathcal{P}_2$ denote the set of $p(u,x|s)$ satisfying $I(U;Z)\geq I(U;S)$. For the channel $p(y,z|x,s)=p(y|x,s)p(z|y)$, any $p\in\mathcal{P}_2$ also satisfies $p\in\mathcal{P}_1$. Therefore, using the Proposition 8, we have that if (R_a,R_l) is achievable, then $(R_a,R_l)\in\mathcal{R}_o^3$, where

$$\mathcal{R}_o^3=\bigcup_{p(u,x|s)}(R_a,R_l) \qquad (126)$$

satisfying

$$\begin{aligned} R_a &\leq \min\{H(S),I(X,S;Y)\} && (127)\\ R_l &\geq I(S;Z,U) && (128)\\ 0 &\leq I(U;Y)-I(U;S). && (129)\end{aligned}$$

It remains to show R_a-R_l bound. We argue as follows.

$$\begin{aligned} n(R_a-R_l) &= I(S^n;Y^n)-I(S^n;Z^n) && (130)\\ &= I(X^n,S^n;Y^n)-I(X^n,S^n;Z^n)-(I(X^n;Y^n|S^n)-I(X^n;Z^n|S^n)) && (131)\\ &\overset{(a)}{\leq} I(X^n,S^n;Y^n)-I(X^n,S^n;Z^n) && (132)\\ &\overset{(b)}{=} I(X^n,S^n;Y^n|Z^n) && (133)\\ &\overset{(c)}{\leq} \sum_{i=1}^{n} I(X_i,S_i;Y_i|Z_i) && (134)\\ &= nI(X,S;Y|Z), && (135)\end{aligned}$$

where (a) and (b) are due to the degradedness condition, and (c) follows from the memoryless property of the channel. The result then follows by taking the union over all joint distributions $p(u,x|s)$. ■

APPENDIX F
PROOF OF THE CONVERSE FOR THEOREM 10

Proof: We bound the rate difference as follows.

$$\begin{align}
n(R_a - R_l) &\leq I(S^n;Y^n) - I(S^n;Z^n) \tag{136}\\
&= \sum_{i=1}^{n} I(S^n;Y_i|Y_1^{i-1}) - I(S^n;Z_i|Z_{i+1}^n) \tag{137}\\
&\overset{(a)}{=} \sum_{i=1}^{n} I(S^n;Y_i|Y_1^{i-1},Z_{i+1}^n) + I(Z_{i+1}^n;Y_i|Y_1^{i-1}) - I(Z_{i+1}^n;Y_i|Y_1^{i-1},S^n) \nonumber\\
&\quad - \left[I(S^n;Z_i|Y_1^{i-1},Z_{i+1}^n) + I(Y_1^{i-1};Z_i|Z_{i+1}^n) - I(Y_1^{i-1};Z_i|Z_{i+1}^n,S^n)\right] \tag{138}\\
&\overset{(b)}{=} \sum_{i=1}^{n} I(S^n;Y_i|Y_1^{i-1},Z_{i+1}^n) - I(S^n;Z_i|Y_1^{i-1},Z_{i+1}^n) \tag{139}\\
&= I(S_i;Y_i|Y_1^{i-1},Z_{i+1}^n,S_1^{i-1},S_{i+1}^n) + I(S_1^{i-1},S_{i+1}^n;Y_i|Y_1^{i-1},Z_{i+1}^n) \nonumber\\
&\quad -I(S_i;Z_i|Y_1^{i-1},Z_{i+1}^n,S_1^{i-1},S_{i+1}^n) - I(S_1^{i-1},S_{i+1}^n;Z_i|Y_1^{i-1},Z_{i+1}^n) \tag{140}\\
&\overset{(c)}{\leq} \sum_{i=1}^{n} I(S_i;Y_i|U_i) - I(S_i;Z_i|U_i) \tag{141}\\
&\overset{(d)}{=} n[I(S;Y|U) - I(S;Z|U)], \tag{142}
\end{align}$$

where, in (a), we used the equalities

$$I(S^n;Y_i|Y_1^{i-1}) + I(Z_{i+1}^n;Y_i|Y_1^{i-1},S^n) = I(Z_{i+1}^n;Y_i|Y_1^{i-1}) + I(S^n;Y_i|Y_1^{i-1},Z_{i+1}^n) \tag{143}$$

and

$$I(S^n;Z_i|Z_{i+1}^n) + I(Y_1^{i-1};Z_i|Z_{i+1}^n,S^n) = I(Y_1^{i-1};Z_i|Z_{i+1}^n) + I(S^n;Z_i|Z_{i+1}^n,Y_1^{i-1}); \tag{144}$$

in (b), we used the Csiszar's sum lemma [27] (and a conditional form of it) to obtain the equalities

$$\sum_{i=1}^{n} I(Z_{i+1}^n;Y_i|Y_1^{i-1}) = \sum_{i=1}^{n} I(Y_1^{i-1};Z_i|Z_{i+1}^n) \tag{145}$$

and

$$\sum_{i=1}^{n} I(Z_{i+1}^n;Y_i|Y_1^{i-1},S^n) = \sum_{i=1}^{n} I(Y_1^{i-1};Z_i|Z_{i+1}^n,S^n); \tag{146}$$

in (c), we define $U_i \triangleq (Y_1^{i-1}, Z_{i+1}^n, S_1^{i-1}, S_{i+1}^n)$; and in (d), we obtain the single-letter expression (by defining $U = (U_i, i)$, etc., see, e.g., [36]). Note that, with the definition of U_i in (c), X_i is generated from S^n, and hence from (U_i, S_i). In addition, given (X_i, S_i), Z_i is independent of (U_i, S_i). Thus, $(U,S) \to (X,S) \to Z$ forms a Markov chain, and the upper bound is given by

$$\begin{align}
R_a - R_l &\leq \max_{p(u,x|s)\text{ s.t. }(U,S)\to(X,S)\to Z} I(S;Y|U) - I(S;Z|U) \tag{147}\\
&= \max_{p(x|u^*,s),\ u^*\in\mathcal{U}} I(S;Y|U=u^*) - I(S;Z|U=u^*) \tag{148}\\
&= \max_{p(x|s)} I(S;Y) - I(S;Z), \tag{149}
\end{align}$$

where the equalities follow due to the following: First, the conditional mutual information expression is maximized with a particular input u^* (as randomizing over different u values will not increase the sum $\sum_u (I(S;Y|U=u) - I(S;Z|U=u))\Pr\{U=u\})$) and a probability distribution $p^*(x|u^*,s)$. Second, the optimal $p^*(x|u^*,s)$ will correspond to a $p(x|s)$. Thus, the converse result can be stated over input distributions in the form $p(x|s)$, thus matching the achievability result. ■

References

[1] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust key generation from signal envelopes in wireless networks," in *Proc. 14th ACM Conference on Computer and Communications Security (CCS)*, 2007.

[2] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepaty extracting a secret key from an unauthenticated wireless channel," in *ACM MOBICOM*, 2008.

[3] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *ACM MOBICOM*, 2009.

[4] S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-theoretically secret key generation for fading wireless channels," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 240–254, Jun. 2010.

[5] H. Delfs and H. Knebl, *Introduction to Cryptography: Principles and Applications*, 2nd ed. Springer, 2007.

[6] O. Goldreich, *Foundations of Cryptography: Volume II, Basic Applications*. Cambridge University Press, 2004.

[7] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: Making software radios more personal," *IEEE Personal Commun. Mag.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[8] J. Mitola III, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Computer Communication System Laboratory, Department of Teleinformatics, Royal Institute of Technology (KTH), Stockholm, Sweden, May 2000.

[9] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[10] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai, and S. Verdu, "Capacity of cognitive interference channels with and without secrecy," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 604–619, Feb. 2009.

[11] O. Simeone and A. Yener, "The cognitive multiple access wire-tap channel," in *43rd Annual Conference on Information Sciences and Systems (CISS 2009)*, Mar. 2009.

[12] L. Toher, O. O. Koyluoglu, and H. E. Gamal, "Secrecy games over the cognitive channel," in *Proc. 2010 IEEE International Symposium on Information Theory (ISIT 2010)*, Austin, TX, Jun. 2010.

[13] J. Zhang and M. Gursoy, "Secure relay beamforming over cognitive radio channels," in *Proc. 45th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2011.

[14] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005–4019, Sep. 2008.

[15] R. Tandon, S. Ulukus, and K. Ramchandran, "Secure source coding with a helper," in *Proc. 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton 2009)*, 2009.

[16] O. O. Koyluoglu and H. El Gamal, "Cooperative encoding for secrecy in interference channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5682–5694, Sep. 2011.

[17] M. Yuksel, X. Liu, and E. Erkip, "A secure communication game with a relay helping the eavesdropper," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 818–830, Sep. 2011.

[18] S. Gel'fand and M. Pinsker, "Coding for channels with random parameters," *Probl. Contr. and Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[19] A. Khisti, S. N. Diggavi, and G. Wornell, "Secret-key generation with correlated sources and noisy channels," in *Proc. 2008 IEEE International Symposium on Information Theory (ISIT'08)*, Jul. 2008.

[20] V. Prabhakaran, K. Eswaran, and K. Ramchandran, "Secrecy via sources and channels: A secret key - Secret message rate tradeoff region," in *Proc. 2008 IEEE International Symposium on Information Theory (ISIT'08)*, Jul. 2008.

[21] A. Khisti, "Secret key agreement on wiretap channels with transmitter side information," in *Proc. 16th European Wireless Conference (EW 2010)*, Lucca, Italy, Apr. 2010.

[22] A. Khisti, S. Diggavi, and G. Wornell, "Secret-key agreement with channel state information at the transmitter," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 672 –681, Sep. 2011.

[23] Y. Chen and A. J. H. Vinck, "Wiretap channel with side information," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 395–402, Jan. 2008.

[24] C. Mitrpant, A. J. H. Vinck, and Y. Luo, "An achievable region for the Gaussian wiretap channel with side information," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2181–2190, May 2006.

[25] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[26] A. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[27] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[28] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1486–1495, Apr. 2005.

[29] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.

[30] N. Merhav and S. Shamai, "Information rates subject to state masking," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2254–2261, Jun. 2007.

[31] J. Körner and K. Marton, "A source network problem involving the comparison of two channels ii," *Trans. Colloquim Inform. Theory*, Aug. 1975, keszthely, Hungary.

[32] ——, "Comparison of two noisy channels," *Topics in Information Theory (Second Colloq., Keszthely, 1975). Amsterdam: North-Holland*, pp. 411–423, 1977.

[33] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1 – 10, Jan. 1976.

[34] C. Heegard and A. E. Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inf. Theory*, vol. 29, no. 5, pp. 731–739, Sep. 1983.

[35] A. V. Kuznetsov and B. S. Tsybakov, "Coding in a memory with defective cells," *Probl. Pered. Inform.*, vol. 10, no. 2, pp. 52–60, 1974.

[36] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley and Sons, Inc., 1991.

[37] C. Heegard, "Capacity and coding for computer memory with defects," Ph.D. dissertation, Stanford University, Stanford, CA, Nov. 1981.